

What Is Pulse Width Modulation

Kevin Craig, Professor of Mechanical Engineering, Marquette University

As engineers, there are many components and devices we use every day that we treat as black boxes. We only are concerned about the inputs and outputs of the black box, and this is acceptable provided we know the operating range and limitations of the device. The same can be said about certain concepts and techniques. We routinely use them and they work, but maybe we really do not understand why they work. As fundamental performance limitations and safety come into question, a deeper understanding becomes necessary. In addition, innovation demands understanding. The widely used pulse-width-modulation (PWM) technique may fall into this category. Most engineers use it every day to drive a variety of devices, but only a few know why it is used and how it works.

A pulse signal is defined by its amplitude and pulse width. A periodic pulse train has a frequency (pulse repetition rate) and a duty cycle (ratio of pulse width to repetition period, varying between 0 percent and 100 percent). In PWM, we modulate the duty cycle while keeping the period fixed. A PWM signal can be generated by microcontrollers, which operate in the digital domain. While an analog signal is continuous in both time and amplitude, a digital signal is discrete in time (sampled at a certain rate) and quantized in amplitude (using a finite number of bits). The output of a microcontroller is typically either digital or PWM. The PWM signal typically varies between 5V and 0V, high and low, and thus can be used to turn an electronic power switch (a transistor) on and off, and control the amount of power delivered to a load.

Why would anybody be interested in this type of signal? PWM avoids losses normally incurred when a power source is limited by resistive means. In PWM, the average power delivered is proportional to the modulation duty cycle. Thus, the switched circuits, used to control the voltage across or current through a load, have low power loss as the switching devices are either off (no current, no power loss) or on (low power loss due to low voltage drop). The PWM signal, with a sufficiently high modulation frequency, can be augmented with a low-pass filter to smooth the pulse train and recover an average analog waveform. PWM has been called a poor man's digital-to-analog (D/A) converter.

What should the frequency of the PWM signal be? First, consider the case where we are using the PWM signal as a D/A converter. A wide variety of microcontroller applications exist that need analog output, but do not require high-resolution D/A converters. In a typical PWM signal, the frequency is constant, but the pulse width (duty cycle) is a variable, directly proportional to the amplitude of the original unmodulated signal. The bandwidth of the low-pass filter should be equal to the bandwidth of the unmodulated signal. The PWM frequency is chosen to give an acceptable ripple magnitude in the analog signal. For example, if a RC low-pass filter is used, the amplitude attenuation in decibels (dB) is given by:

$$\text{dB} = 20 \log_{10} \frac{1}{\sqrt{\left(f_{\text{PWM}} / f_{\text{BW}}\right)^2 + 1}}$$

To increase the attenuation, and thus reduce the ripple, a higher-order filter may be necessary or a higher PWM frequency.

Many devices inherently average an on-off signal to control their operation, based on the duty cycle. Examples include LEDs viewed by humans and inductive loads, such as motors and solenoids. For an inductive load, e.g., a LR circuit, what should the PWM voltage frequency be so that the current waveform is within a certain percentage (P percent) of the analog step response? A fundamental analysis of a LR circuit shows that frequency (f Hz) to be:

$$f = - \frac{R}{2L \ln \left(1 - \frac{P}{100} \right)}$$

Too often, devices and concepts are shrouded in mystery, which can often lead to avoidance or misuse. Focusing on the fundamentals removes that mystery.